CARNEGIE INSTITUTE OF TECHNOLOGY

DEPARTMENT OF MATHEMATICS

Pittsburgh 13, Pennsylvania

CHARACTERIZING NORMAL LAW AND A NONLINEAR INTEGRAL EQUATION

M. M. Rao

TECHNICAL REPORT NO. 7

prepared under

CONTRACT NO. DA-36-061-ORD 447

Department of the Army Project No. 5B99-01-004
Ordnance Research and Development No. TB2-0001
Office of Ordnance Research Project No. 1280

CIT-ORD-16-TR7

July, 1962

CARNEGIE INSTITUTE OF TECHNOLOGY

DEPARTMENT OF MATHEMATICS

Pittsburgh 13, Pennsylvania

CHARACTERIZING NORMAL LAW AND A NONLINEAR INTEGRAL EQUATION

M. M. Rao

TECHNICAL REPORT NO. 7

prepared under

CONTRACT NO. DA-36-061-ORD 447

Department of the Army Project No. 5B99-01-004
Ordnance Research and Development No. TB2-0001
Office of Ordnance Research Project No. 1280

CIT-ORD-16-TR7

July, 1962

ABSTRACT

The problem considered in this note is to characterize the distribution function $F(x)$, given that the mean and a "smooth tube statistic" (e.g. sample range, variance etc.) formed of a random sample from $F(x)$ are statistically independent. The problem leads to a nonlinear integral equation. Under some mild restrictions on the class of $F(x)$, such as continuity and the behavior in the tails, the solution of the integral equation was found to be the normal law and this gives the characterization of $F(x)$.

The study of the nonlinear integral equation presented here may be of independent interest since very little seems to be known about such equations in the literature. The result also implies that the means and range in random samples will be independent only for normal distributions, and this is important for "chart procedures" in Quality Control practice.

CHARACTERIZING NORMAL LAW AND A NONLINEAR INTEGRAL EQUATION

M. M. Rao [1]

1. Introduction: Let $X_1, \ldots, X_n$ be independent random variables (r.v.'s) with common distribution function (d.f.) $F(x)$. If two different functions of these r.v.'s (usually one linear and the other non-linear) are given to be statistically independent, then in a number of cases the d.f. $F(x)$ was characterized completely. For instance if these functions are the mean and the variance, then $F(x)$ was normal. (See Lukacs [6].) A general method is to translate the conditions of independence, usually through Fourier transforms, into a differential equation whose solution then yields the desired result. Certain other related methods were also discussed in [6]. However, the use of integral equations in these problems did not seem to have been considered in the past. The purpose of this paper is to characterize the normal distribution through a nonlinear integral equation.

The problem considered in this paper is to characterize the d.f. $F(x)$, given that the mean and a "tube statistic" (to be defined below), formed of the above r.v.'s, are independently distributed. The so-called tube statistics include "sample range", "sample central absolute moments", etc. In the present set up the problem leads naturally to a non-linear integral equation. Under some mild restrictions on the class of d.f.'s $F(x)$, such as continuity and behavior in the tails, the solution of the integral equation was found to be normal, thus giving the characterization of $F(x)$. On the other hand, if $F(x)$ is three times differentiable, the problem was solved in a different way (reducing to a differential equation) by Paskevich [7]. The solution of the integral equation presented here may

---

[1] This work was done in part under Contract No. DA-36-061-ORD-447, and under grant NSF-G 14832. Reproduction in whole or in part is permitted for any purpose of the United States Government.

itself be of independent interest since very little seems to be known about such equations in the literature.

Before proceding to the next section, it is of interest to note a consequence of the result given in Sections 5 and 6. It implies that the mean and range in random samples will be independent only for the normal d.f., so that in Quality Control practice the "chart procedures" must be limited to the normal distributions.

2. Tubular Functions: Let $(x_1,\ldots,x_n)$ be an n-vector of real numbers and let $h(x_1,\ldots,x_n)$ be a real valued non-negative continuous function of $(x_1,\ldots,x_n)$. Then h is said to be a tubular function if the following three conditions hold:

(i) $h(x_1,\ldots,x_n) = 0$ if, and only if, $x_1 = x_2 = \cdots = x_n$.

(ii) $h(x_1,\ldots,x_n) = \alpha$ defines a (not necessarily circular) cylinder with $x_1 = \cdots = x_n$ as the axis. [This implies $h(x_1,\ldots,x_n) = h(x_1 + c,\ldots,x_n + c)$ for all $x_1,\ldots,x_n$ and c.]

(iii) Corresponding to any pair of numbers a, b, in the range of h, there exists a constant c (depending on a and b) such that $h(x_1,\ldots,x_n) = a$ if, and only if, $h(cx_1 + (1 - c)\bar{x},\ldots, cx_n + (1 - c)\bar{x}) = b$, where $\bar{x} = \sum_{i=1}^{n} x_i/n$.

Thus a tubular function is translation invariant and for any two numbers a_1, a_2 in the range of h, $h(x_1,\ldots,x_n) = a_i$, $i = 1, 2$, determine similar (or homothetic) cylinders. The above definition can be given a parametric form, as follows, [7]: (i.e. the surface $h(x_1,\ldots,x_n) = \alpha$ can be written as (for $n \geq 3$))

$$
(1) \qquad \begin{aligned}
x_1 &= t + \phi(\alpha) \quad \psi_1(\theta_1,\ldots,\theta_{n-2}) \\
x_2 &= t + \phi(\alpha) \quad \psi_2(\theta_1,\ldots,\theta_{n-2}) \\
&\vdots \\
x_n &= t + \phi(\alpha) \quad \psi_n(\theta_1,\ldots,\theta_{n-2})
\end{aligned}
$$

where $-\infty < t < \infty$, $0 \leq \Theta_j \leq 2\pi$, $j = 1,\ldots,n-2$, $\phi(\alpha) \geq 0$, and $\sum_{i=1}^{n} \psi_i(\underset{\sim}{\Theta}) = 0$, and where there exists an α_o such that $\phi(\alpha_o) = 0$. This is a familiar representation used in statistics with $\psi_i(\Theta_1,\ldots,\Theta_{n-2})$ as a product of trigonometric functions.

In what follows a slightly restricted class of h's will be considered, but actually the procedures presented below, with some modifications, are applicable to a wider class.

<u>Definition</u> 1: A tubular function is said to be <u>smooth</u> if the parametric form satisfies the following conditions also:

(i) The functions $\phi(\alpha)$, $\psi_i(\underset{\sim}{\Theta})$, $0 \leq \Theta_j \leq 2\pi$, have continuous partial derivatives in $(0,\infty)$ and in (the cartesian product) $\bigtimes_{j=1}^{n-2} (0,2\pi)$ respectively, so that the Jacobian D of the transformation exists.

(ii) $D \neq 0$ for some point $(t,\alpha,\Theta_1,\ldots,\Theta_{n-2})$ of the transformed space. [Note that D does not depend on t. In fact, $D = \phi^{n-2}(a)\ \phi'(\alpha)\tilde{D}(\underset{\sim}{\Theta})$.]

(iii) $\int_0^{2\pi} \cdots \int_0^{2\pi} |\tilde{D}(\underset{\sim}{\Theta})| d\Theta_1,\ldots,d\Theta_{n-1}$ exists.

These assumptions are made in order that the inverse transformation may exist in some neighborhood of a point, and calculations may be made through the Inverse Function Theorem.

If $X_1,\ldots,X_n$ are r.v.'s then $h(X_1,\ldots,X_n)$ is called a <u>tube statistic</u> where h is defined above. As examples of such statistics one has

(i) $h(X_1,\ldots,X_n) = \max_{1 \leq i,\, j \leq n} |X_i - X_j|$ (sample range);

(ii) $h(X_1,\ldots,X_n) = \frac{1}{n} \sum_{i=1}^{n} |X_i - \bar{X}|^k$, where $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ and $k \geq 1$ (sample k^{th} absolute moment).

Throughout this paper, it is always assumed that $X_1, \ldots, X_n$ are independent, r.v.'s having a continuous density $g(x)$ defined for all x on the line. Let $\bar{X}$ be the sample mean and T be a tube statistic. Throughout, the tube statistic T is assumed to be smooth as defined above. A consequence of smoothness is that T (as well as $\bar{X}$) has a well-defined density function on the line.

3. <u>A Lemma</u>: In this section a lemma that is to be the starting point of this investigation is proved. It is stated (without proof and less tightly) in the course of an argument in [7]. (See also the remarks below at the end of this section.)

<u>Lemma</u> 1: Let $X_1, \ldots, X_n$ be n independent, and identically distributed r.v.'s each with a continuous density function (relative to the Lebesgue measure) $g(x)$ defined for all x on the line. Let $\bar{X}$ and T be the mean and a smooth tube statistic respectively, based on the above r.v.'s. Let $f(x)$ be the density function of $\bar{X}$. If $\bar{X}$ and T are independent, then it follows that $g^n(x) = af(x)$ a.e., where $a = \int_{-\infty}^{\infty} g^n(x)dx$.

<u>Proof</u>: From the remark at the end of Section 2, T and $\bar{X}$ have densities (relative to Lebesgue measure), and from the classical procedure of transformation of variables, one has the joint and individual densities of T and $\bar{X}$ given by,

$$(2) \quad \frac{\partial^2 \Pr\{T < \alpha, \bar{X} < x\}}{\partial\alpha\, \partial x} = \int_0^{2\pi} \cdots \int_0^{2\pi} \Big[\prod_{i=1}^{n} g(x + \phi(\alpha) \quad \psi_i(\underset{\sim}{\theta}))\Big] |D| . d\theta_1 \ldots d\theta_{n-2}$$

$$(3) \quad \frac{\partial \Pr\{T < \alpha\}}{\partial\alpha} = \int_{-\infty}^{\infty} \frac{\partial^2 \Pr\{T < \alpha, \bar{X} < x\}}{\partial\alpha\, \partial x} dx.$$

By hypothesis T and $\bar{X}$ are independently distributed. For this it is necessary and sufficient that the joint density of T and $\bar{X}$, given by (2),

splits into a function of α alone (the density of T) and a function of x alone (the density of $\bar{X}$) a.e. Since D, the Jacobian, does not involve x (cf. (ii) Definition 1) and can be written as $\phi^{n-2}(\alpha)\ \phi'(\alpha)\tilde{D}(\underset{\sim}{\theta})$, the right side of (2) takes the following form.

$$(4) \quad \int_0^{2\pi} \cdots \int_0^{2\pi} \left[\prod_{i=1}^{n} g(x + \phi(\alpha)\ \psi_i(\underset{\sim}{\theta}))\right] |\tilde{D}(\underset{\sim}{\theta})| d\theta_1, \ldots, d\theta_{n-2} = f(x)\cdot I(\alpha),$$

where $I(\alpha) = \tilde{g}(\alpha)/\phi^{n-2}(\alpha)\ \phi'(\alpha)$, and where $\tilde{g}(\alpha)$ and $f(x)$ are the densities of T and $\bar{X}$. But (4) is an identity in $-\infty < x < \infty$, and all α for which $I(\alpha)$ is well defined. Note that if $\phi(\alpha) = 0$, for some α, then $\tilde{g}(\alpha) = 0$ also for the same α. By hypothesis there exists an α_0 such that $\phi(\alpha_0) = 0$. Thus taking $\alpha \longrightarrow \alpha_0$, through a sequence of α's in the domain of $\phi(\alpha)$, in (4) and remembering the conditions (ii) and (iii) of the definition of smooth tube statistic, it follows by the Dominated Convergence Theorem, that the limit on the left side of (4) can be taken inside the integral, and that the limit exists ($\neq 0$), so that on the right side $\lim_{\alpha \to \alpha_0} I(\alpha) = c_0$, ($> 0$) also exists. Since the left-side becomes $g^n(x) \cdot c'$ (c' is a positive constant), the lemma follows.

<u>Remarks</u>: The condition of smoothness can be dropped and other conditions (less stringent) can be imposed using another procedure as follows. From the definition of tube statistic and the continuity of the d.f. of X_i's, one has, for any $\delta > 0$,

$$(5) \qquad 0 < \Pr\{T < \delta\} \quad \text{and} \quad \Pr\{T < \delta\} \longrightarrow 0, \text{ as } \delta \longrightarrow 0.$$

If $t_1 < t_2$ are two arbitrarily fixed real numbers, then from the independence assumption of $\bar{X}$ and T it results that

$$(6) \qquad \Pr\{t_1 < \bar{X} < t_2\} = \frac{\Pr\{t_1 < \bar{X} < t_2,\ T < \delta\}}{\Pr\{T > \delta\}}.$$

Since $\delta > 0$ is arbitrary, one can take $\delta \longrightarrow 0$ on both sides and, since the left side is independent of δ, the limit on the right exists. The problem then is to calculate the limit. This is not entirely trivial, and a suitable procedure is a method using Fourier analysis analogous to the one found in [5]. For this it suffices to assume that the event $\{T < \delta\}$ can be approximated, for small $\delta > 0$, by the event $\{|X_i - \bar{X}| < \varepsilon_i,\ i = 1,\ldots,n\}$. The lemma was originally proved in this way, but the calculations are long and tedious. It is useful to note that under these different sets of assumptions on T, one has the same conclusion as is given in the above lemma.

4. A Nonlinear Integral Equation: The conclusion of Lemma 1 can be rewritten as follows. Since $\bar{X} = \sum_{i=1}^{n} (X_i/n)$, the density function of $\bar{X}$ is the n-fold convolution of the densities of (X_i/n). The latter is $ng(nx)$. Since also $af(x) = g^n(x)$, which is the conclusion of the lemma, and $f(x)$ is the density of $\bar{X}$ the following equation obtains.

$$(7) \qquad g^n(x) = na \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(nx - u_1 - \cdots - u_{n-1}) g(u_1) \cdots g(u_{n-1}) du_1 \cdots du_{n-1},$$

where $a = \int_{-\infty}^{\infty} g^n(x)dx$. Now (7) is a non-linear integral equation whose solutions $g(x)$ subject to (i) $g(x) \geq 0$, (ii) $\int_{-\infty}^{\infty} g(x)dx = 1$ are the required ones. Since the equation for general n does not essentially involve any new difficulties, it suffices to consider the case $n = 2$. Thus (7) can be written as (see also the parenthetical remark after (11) below about the case $n > 3$)

$$(8) \qquad g^2(x) = 2a \int_{-\infty}^{\infty} g(2x - u)\, g(u)\, du, \qquad a = \int_{-\infty}^{\infty} g^2(x)dx.$$

It is required to consider the equation (8) for $g(x)$ satisfying (i) and (ii) above, and no other conditions being imposed. Unfortunately there

are no general techniques available to solve non-linear integral equations, even those of type (8). So it is considered here in detail, and this study may be of independent interest.

Consider the function $g^*(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\theta)^2\right\}$ for $-\infty < \theta < \infty$, $0 < \sigma^2 < \infty$. It is readily verified that this function for each θ and σ^2 is a solution of (8). This is only a particular solution. To determine all solutions of (8), satisfying (i) and (ii), define the function

$$(9) \qquad h(x) = g(x)/g^*(x).$$

Since $g^*(x) > 0$ for all x, h(x) is well defined for all x on the line. Substituting (9) into (8), one gets,

$$(10) \qquad h^2(x) = 2a \int_{-\infty}^{\infty} e^{-v^2/\sigma^2} h(x+v)\, h(x-v)\,dv,$$

$$(11) \qquad a = \frac{1}{2\pi\sigma^2} \int_{-\infty}^{\infty} h^2(u)\, e^{-(u-\theta)^2/\sigma^2}\,du.$$

(If $n > 2$, one gets a positive definite quadratic form of v's in the exponent of the exponential in (10) above instead of v^2. Otherwise the procedure is the same.) From (8) and (10), some information about g(x), and h(x) is obtained:

<u>Proposition</u> 1: Every non-negative measurable solution g(x) of the integral equation (8) belongs to $L^p(R,\mu)$, $1 \leq p \leq \infty$, where $L^p(R,\mu)$ is the Lebesgue space on the line R with μ as the ordinary Lebesgue measure. Moreover, for all $x \in R$, $g(x) > 0$. (As usual, $d\mu$ and dx are interchanged below.)

<u>Remark</u>: If g(x) is also assumed to be continuous, this is immediate. But it is of interest to study (8) generally.

Proof: Equation (8) can be written as

$$(12) \qquad \int_{-\infty}^{\infty} g(z-u)\, g(u)du = w(z) \qquad \left(= \frac{1}{2a} g^2(x)\right).$$

Thus $w(z)$ being the convolution of two $L^1(R,\mu)$ functions is itself in $L^1(R,\mu)$ and is defined for all $z\epsilon R$. Moreover, since $w(z) = w(2x)$,

$$(13) \qquad 1 = \int_{-\infty}^{\infty} w(z)dz = 2\int_{-\infty}^{\infty} w(2x)dx = \frac{1}{a}\int_{-\infty}^{\infty} g^2(x)dx,$$

shows that $g(x)\epsilon L^2(R,\mu)$. According to a theorem in Linear Analysis (cf., e.g. [2], p. 528) $w(z)\epsilon L^s(R,\mu)$ where $s^{-1} = \frac{1}{2} + \frac{1}{2} - 1$, i.e. $s = +\infty$, using the fact (just established) that $g(x)\epsilon L^2(R,\mu)$. Thus every solution of (8) is an essentially bounded function. From this it follows that, if $1 < \alpha < \infty$,

$$(14) \qquad \int_{-\infty}^{\infty} g^{\alpha}(x)dx = \int_{-\infty}^{\infty} g^{\alpha-1}(x)d\tilde{\lambda}(x) < \infty,$$

where $\tilde{\lambda}$ is the finite measure given by $d\tilde{\lambda} = g(x)dx$. The first part of the proposition is proved.

Now, for the second part, if possible let $x = x_o$ be such that $g(x_o) = 0$. Since $g^*(x_o) > 0$, this is equivalent to (cf. (9)), $h(x_o) = 0$. But from (10), since $h(x) \geq 0$, it follows that $h(x_o \pm v) = 0$, a.e. $[\mu]$. Then this implies that $g(x) = 0$ a.e. $[\mu]$, which is impossible, since $\int_{-\infty}^{\infty} g(x)dx = 1$. This completes the proof of the proposition.

The above proof has, as a consequence, the following

Corollary: The function $h(x)$ of (9) is strictly positive for all real x,

and is bounded on every gounded interval of the line R.

The next result is an essential step toward the solution of the integral equation (8). It is the following:

Proposition 2: If h(x) defined by (9) satisfies the condition

$$\lim_{x \to \infty} h(x) = c_1, \quad \lim_{x \to -\infty} h(x) = c_2, \quad 0 \leq c_1, c_2 < \infty, \text{ or } \quad 0 < c_1, c_2 \leq \infty, \tag{15}$$

then $h(x) = \alpha$, a constant, for all $x \varepsilon R$, so that, under (15), this is the only solution of (10).

Proof: First it will be shown that $0 < c_i < \infty$. Suppose the contrary. Consider the case $c_i = 0$ for at least one i $(= 1,2)$. Let $f_x(v) = h(x + v)\, h(x - v)$.

Rewrite (10) in a convenient form as

$$h^2(x) = \int_{-\infty}^{\infty} f_x(v) d\lambda(v) \tag{16}$$

where for every Borel set E on the line R,

$$\lambda(E) = 2a \int_E e^{-v^2/\sigma^2} dv, \tag{17}$$

so that λ is a finite measure, and is equivalent to the Lebesgue measure (i.e., $\lambda \equiv \mu$). Thus $f_x(v) \varepsilon L^1(R, \lambda)$ for every $x \varepsilon R$.

Now let S be a transformation on R such that $Sv = v + 1$. Then S and S^{-1} are measurable transformations (relative to λ or μ), with the following properties relative to λ :

(i) if $E \subset R$ is a Borel set, $SE \subset E$ implies $\lambda(E - SE) = 0$, i.e., S is incompressible. This is immediate since S only translates.

(ii) if E is as in (i), $\lambda(SE) = 0$ and $\lambda(S^{-1}E) = 0$, if, and only if, $\lambda(E) = 0$. This is also immediate, since $\lambda \equiv \mu$, $\lambda(SE) = 0$

implies $\mu(SE) = \mu(E) = 0$, and then $\lambda(E) = 0$. Similarly for $\lambda(S^{-1}E)$. (i.e. S is non-singular.) However, S is not λ-measure preserving.

It will now be shown that the assumption $c_i = 0$ leads to a contradiction. This is accomplished, with the application of the above transformation S to $f_x(v)$, through Hurewicz's ergodic Theorem [4], in Halmos' form [3].

Now consider, for every Borel set $E \subset R$,

$$(18) \qquad \nu(E) = \int_E f_x(v)\, d\lambda(v).$$

Since $f_x(v) > 0$ by the preceding corollary, $\nu \equiv \lambda$. Let

$$(19) \qquad \nu_n(E) = \sum_{i=0}^{n-1} \nu(S^i E), \quad \lambda_n(E) = \sum_{i=0}^{n-1} \lambda(S^i E), \quad S^o = I$$

where I is the identity transformation. Consequently,

$$(20) \qquad \nu_n(E) = \int_E \sum_{i=0}^{n-1} f_x(S^i v) d\lambda(S^i v), \qquad \lambda_n(E) = \sum_{i=0}^{n-1} \int_E d\lambda(S^i v).$$

Since $\nu_n << \lambda_n$, by the Radon-Nikodym (R-N) Theorem, there exists a finite function $k_x^{(n)}(v)$ (≥ 0) such that

$$(21) \qquad \nu_n(E) = \int_E k_x^{(n)}(v) d\lambda_n(v) = \sum_{i=0}^{n-1} \int_{S^i E} f_x(v)\, d\lambda(v),$$

for every measurable set $E \subset R$. Now by Hurewicz's theorem [4],

$$(22) \qquad \lim_{n \to \infty} k_x^{(n)}(v) = k_x^*(v) \quad \text{a.e. } [\lambda],$$

exists and $k_x^*(Sv) = k_x^*(v)$, i.e., k_x^* is invariant. Moreover, since λ is a

finite measure, one also has [4],

$$(23)\qquad h^2(x) = \int_{-\infty}^{\infty} f_x(v)\,d\lambda(v) = \int_{-\infty}^{\infty} k_x^*(v)\,d\lambda(v).$$

For the present proof, however, it is necessary to compute $k_x^*(v)$ explicitly, and for this $k_x^{(n)}(v)$ is also needed. The calculation is as follows: Since S is measurable and non-singular, $\lambda S^i << \lambda$, for any i. So by the R-N theorem, there exists a non-negative finite valued measurable function $w_i(v)$, which may be taken to be positive, such that for every measurable set E,

$$(24)\qquad \lambda(S^i E) = \int_E w_i(v)\,d\lambda(v).$$

Using the properties of R-N derivatives, from (19) and (24) one gets,

$$(25)\qquad \nu_n(E) = \int_E \sum_{i=0}^{n-1} f_n(S^i v) w_i(v)\,d\lambda(v), \quad \lambda_n(E) = \int_E \sum_{i=0}^{n-1} w_i(v)\,d\lambda(v).$$

From (21) and (25), since $\lambda_n \equiv \lambda$, the following result obtains.

$$(26)\qquad \nu_n(E) = \int_E k_x^{(n)}(v)\,d\lambda_n(v) = \int_E \frac{\sum_{i=0}^{n-1} f_x(S^i v)\, w_i(v)}{\sum_{i=0}^{n-1} w_i(v)}\,d\lambda_n(v).$$

Since E is an arbitrary measurable set one deduces from (26),

$$(27)\qquad k_x^{(n)}(v) = \frac{\sum_{i=0}^{n-1} f_x(S^i v)\, w_i(v)}{\sum_{i=0}^{n-1} w_i(v)}, \quad \text{a.e. } [\lambda],$$

where $k_x^{(n)}(v) \longrightarrow k_x^*(v) = k_x^*(Sv)$, a.e. $[\lambda]$. But by definition, $f_x(S^n v) =$

$h(x + v + n)\, h(x - v - n)$. Because of the form of $f_x(v)$ it follows, from $k_x^*(Sv) = k_x^*(v)$, that k_x^* is independent of v, but possibly a function of x. Also by (15) (after noting the form of $f_x(S^n v)$), and the assumption that $c_i = 0$, one concludes that $\lim_{n \to \infty} f_x(S^n v) = 0$.

Now $w_i(v) > 0$ and, since S is non-singular and incompressible, $\lim_{n \to \infty} \sum_{i=0}^{n-1} w_i(v) = +\infty$, for almost all v, by Theorem 3 of Halmos [3]. Then setting $a_{ni} = w_i(v) / \sum_{i=0}^{n-1} w_i(v)$, it follows that (i) $a_{ni} \longrightarrow 0$ as $n \longrightarrow \infty$ for every $i = 0,1,\ldots$, and (ii) $\sum_{i=0}^{n-1} a_{ni} = 1$, so that a_{ni} satisfy, for almost all v, the Toeplitz conditions in sumability theory. So $k_x^{(n)}(v) = \sum_{i=0}^{n-1} a_{ni} f_x(S^i v)$ is 'summable'. More precisely,

$$\text{(28)} \qquad \underline{\lim}_{n \to \infty} f_x(S^n v) \leq \underline{\lim}_{n \to \infty} k_x^{(n)}(v) \leq \overline{\lim}_{n \to \infty} k_x^{(n)}(v) \leq \overline{\lim}_{n \to \infty} f_x(S^n v).$$

(Cf. Zygmund [8], p. 75, Theorem 1.3.) Since the extremes are zero by the hypothesis and supposition, and the middle terms equal k_x^* by the preceding analysis, it follows that $k_x^* = 0$, for every real x. By (23) then $h(x) = 0$ for every x. This contradicts the corollary to Proposition 1, and hence $c_i > 0$ for $i = 1,2$.

Next suppose $c_i = +\infty$ in (15) for at least one i $(= 1,2)$. This gives $\lim_{n \to \infty} f_x(S^n v) = +\infty$ for every x (since $c_i > 0$). Then by (28) which is valid for this case also (cf. [8], p. 75), $k_x^* = +\infty$. This, after using (23), again contradicts the same proposition.[2] Hence $0 < c_i < \infty$. Letting (cf. (15)) $\lim_{n \to \infty} f_x(S^n v) = c$, which exists by the preceding analysis ($c = c_1 c_2$ a constant)

---

[2] Alternately, this also contradicts Hurewicz's theorem, [4], according to which k_x^* is finite a.e.

one has, by (28), $k_x^* = c$, a constant.

Now from what precedes it results that (cf. (23))

$$(29) \qquad h^2(x) = \int_{-\infty}^{\infty} c\, d\lambda(v) = 2a.c. \int_{-\infty}^{\infty} e^{-v^2/\sigma^2}\, dv = \alpha^2,$$

say, where α does not depend on x. Thus $h(x) = \alpha$, as was to be proved.

Remarks: This proof also shows that (15) can be replaced by $\lim_{n \to \infty} h(x_o - kx) \prod_{i=1}^{k} h(x_i + x) = c_k$, $0 \leq c_k \leq \infty$, exists for each k. "Simple approaches" do not seem to yield this intuitively obvious result.

It is now possible to state the main result of this section as

Theorem 1: Let $g(x)$ be a non-negative measurable function on the line such that

$$\text{(i)} \quad g^2(x) = 2a \int_{-\infty}^{\infty} g(2x - u)g(u)du, \quad a = \int_{-\infty}^{\infty} g^2(x)dx, \quad \int_{-\infty}^{\infty} g(x)\, dx = 1,$$

$$\text{(ii)} \quad \lim_{x \to \infty} e^{\frac{(x-\theta)^2}{2\sigma^2}} g(x) = c', \quad \lim_{x \to -\infty} e^{\frac{(x-\theta)^2}{2\sigma^2}} g(x) = c'', \quad 0 \leq c', c'' < \infty, \text{ or}$$
$$0 < c', c'' \leq \infty,$$

for some constant $\sigma^2 > 0$, and some real θ. Then all non-negative solutions of the non-linear integral equation (i) are the following (θ, σ^2 being as in (ii)):

$$(30) \qquad g(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2\sigma^2}(x - \theta)^2 \right\}.$$

Proof: It suffices to note that condition (ii) of the theorem on $g(x)$ is precisely the condition (15) for $h(x)$, and then, by Proposition 2, $h(x) = \alpha$, a constant. One finds on substitution in (10) and (11) that $\alpha = 1$. In view of (9), this completes the proof of the theorem.

Remark: It may be that Proposition 2 and (hence) Theorem 1, are true without the Conditions (15) and (ii) above. But the method of proof given here does

not work. This is seen from (23), since (28) does not give any new information. The only conclusion is that k_x^* is a constant multiple of $h^2(x)$ and so $h(x)$ cannot be determined by the present procedure.

In the next section the characterization problem of the probability law, stated in the introduction, is considered.

5. The Characterization Problem: The characterization of the d.f. of X_i's when the mean and tube statistic are independently distributed can now be given. This is the main result and is contained in

Theorem 2: Let $X_1,\ldots,X_n$ be n (≥ 3) independent r.v.'s with a common d.f. having a continuous density (relative to the Lebesgue measure) $g(x)$. If $\bar{X}$ is the mean of the r.v.'s and $T(X_1,\ldots,X_n)$ is a (smooth) tube statistic and if

$$(*)\qquad \lim_{x\to\infty} e^{\frac{(x-\theta)^2}{2\sigma^2}} g(x) = c', \quad \lim_{x\to-\infty} e^{\frac{(x-\theta)^2}{2\sigma^2}} g(x) = c'', \quad 0 \leq c',c'' < \infty, \text{ or } 0 < c',c'' \leq \infty,$$

for some $\sigma^2>0, \theta\in R$, then $\bar{X}$ and T are independently distributed if, and only if, $g(x)$ is given by (same θ and σ^2 as in (*) above)

$$(30)\qquad g(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\theta)^2\right\},$$

Proof: If $g(x)$ is given by (30), then (*) is automatically satisfied, and it is long known that $\bar{X}$ and T are independently distributed (cf., e.g. [1]), and indeed the proof of the result, using the fact that T can be written as a function of $(X_i - X_j)$, (translation invariance) is quite easy.

The "only if" is hard. Under the conditions of the theorem, by Lemma 1, $g(x)$ satisfies the integral equation (7), and under the remaining condition (*), the conclusion is an immediate consequence of the preceding theorem.

Remark: Just as in the case of Theorem 1, the above theorem may hold without (*), which is a condition about the behavior of $g(x)$ at the tails.

Under different conditions on $g(x)$, involving the differentiability requirements, the result analogous to that of Theorem 2 was given by Paskevich in [7]. Since the method there is entirely different from the one presented here, and since it has some interest of its own, and also because of the inaccessibility of the paper, that result will be given below with a (slightly completed) proof.

6. Another Characterization: In [7], Paskevich gave another useful procedure toward the solution of the characterization problem. The conditions on $g(x)$ given in Theorem 2 above may hold when the corresponding conditions given in [7] fail. For instance, $g(x) = \frac{1}{2} e^{-|x|}$. On the other hand no such growth condition, as (*) of Theorem 2, is assumed in [7].

The main result of (7) is given in the following.

Theorem 3: (Paskevich) Let $X_1,\ldots,X_n$ be independent r.v.'s ($n \geq 3$), with a common d.f. whose density function $g(x)$ is twice differentiable on the line R. Suppose $\bar{X}$ is the mean and T, a (smooth) tube statistic based on the X_i above. Then $\bar{X}$ and T are independently distributed if, and only if, $g(x)$ is normal. (i.e., the density given by (30).)

Proof: As in the preceding theorem, the "if" part is long known, [1], and only the converse need be considered.

"Only if" Part. It is sufficient to consider the case $n = 3$. The differentiability conditions on $g(x)$ lead now to a differential equation, instead of the integral equation of the preceding sections. Since $\bar{X}$ and T are independent r.v.'s (by hypothesis now), one obtains as before,

$$(31) \qquad \frac{\partial^2 \Pr\{T < \alpha, \bar{X} < x\}}{\partial\alpha\, \partial x} = \frac{\partial \Pr\{T < \alpha\}}{\partial\alpha} \cdot \frac{\partial \Pr\{\bar{X} < x\}}{\partial x},$$

Substituting the relevant expressions from (2) and (3) in (31) one gets on simplification, (n = 3 now)

$$(32)\quad \int_0^{2\pi} [\prod_{i=1}^{3} g(x + \phi(\alpha)\Psi_i(\Theta))]|\tilde{D}|d\Theta = f(x)\int_{-\infty}^{\infty}\int_0^{2\pi}\prod_{i=1}^{3} g[t + \phi(\alpha)\Psi_i(\Theta)]|\tilde{D}|d\Theta dt.$$

Letting $A = \phi(\alpha)$, and differentiating (32) relative to A, it becomes (recall that $\tilde{D}$ depends only on Θ)

$$\int_0^{2\pi} [g'(x + A\Psi_1)\ \Psi_1 g(x + A\Psi_2)\ g(x + A\Psi_3) +$$

$$g(x + A\Psi_1)\ g'(x + A\Psi_2)\ \Psi_2 g(x + A\Psi_3) +$$

$$(33)\quad g(x + A\Psi_1)\ g(x + A\Psi_2)\ g'(x + A\Psi_3)\Psi_3]|\tilde{D}|d\Theta$$

$$= f(x)\cdot\int_{-\infty}^{\infty}\int_0^{2\pi} [g'(x + A\Psi_1)\Psi_1 g(x + A\Psi_2)\ g(x + A\Psi_3)+\cdots+\cdots]\,|\tilde{D}|d\Theta dt.$$

Now $\phi(\alpha)$ is a continuous function of α, and for some value $\alpha = \alpha_0, A = \phi(\alpha_0) = 0$. (See (1) and the following.) So in the above, letting A = 0, it reduces to an equation in x. However, nothing new is obtained in (33), since it reduces to the form 0 = 0. For this reason the second derivative of g(x) is assumed to exist, and so differentiating (33) relative to A again, and setting A = 0 in the resulting expression one gets, upon using the condition $\sum_{i=1}^{3}\Psi_i = 0$ and simplifying,

$$(34)\qquad g''(x)\ g^2(x)H - g'^2(x)\ g(x)H = f(x)\cdot E$$

where E and H are some constants. In fact one finds that $H = \int_0^{2\pi}(\Psi_1^2+\Psi_2^2+\Psi_3^2)|\tilde{D}|d\Theta$, so that $H \neq 0$, since the Jacobian in the definition of smooth tube statistic was assumed to be non-vanishing. Thus (34) can be written equivalently as

$$(35) \qquad \frac{d}{dx}\left(\frac{g'(x)}{g(x)}\right) = \frac{f(x)}{g^3(x)} \cdot \frac{E}{H} = B ,$$

where $B = \frac{E}{a \cdot H}$, and where $\frac{1}{a} = \frac{f(x)}{g^3(x)}$ was given by Lemma 1.

The solution of this differential equation (35) satisfying the boundary condition that $g(x)$ be a probability density function is immediately seen to be the normal density. Here it was taken that $B \neq 0$ in (35). If $B = 0$, however, the resulting solution which is a probability density does not satisfy the relation $g^n(x) = af(x)$ which must be true whenever $\bar{\bar{X}}$ and T are independent (Lemma 1). This establishes the converse part and with it the theorem.

The author is indebted to Professor M. H. DeGroot for stimulating conversations during the course of writing this paper. In particular, the material through Proposition 1 constitutes a joint effort.

References

[1] Daley, J. F., "On the use of sample range in an analogue of Student's t-test," Ann. Math. Stat., 17, 71-74, 1946

[2] Dunford, N. and J. T. Schwartz, Linear Operators, Part 1: General Theory. Interscience Publishers, New York, 1958.

[3] Halmos, P. R., "An ergodic theorem," Proc. Nat. Acad. Sci., U.S.A., 32, 156-161, 1946.

[4] Hurewicz, W., "Ergodic theorem without invariant measure," Ann. Math., 45, 192-206, 1944.

[5] Kac, M., "On the characteristic functions of the distributions of estimates of various deviations in samples from a normal population," Ann. Math. Stat., 19, 257-261, 1948.

[6] Lukacs, E., "Characterization of populations by properties of suitable statistics," Proc. Third Berkeley Symp. on Math. Stat. & Prob., 2, 195-214, 1956.

[7] Paskevich, V.S., "On one characteristic of control charts used in quality control," Prikladnaya mat i mekh., 17, 513-516, 1953. (in Russian).

[8] Zygmund, A., Trigonometric Series-I, (2nd edition), Cambridge University Press, London, 1959.

Carnegie Institute of Technology
Pittsburgh, Pennsylvania

OFFICE OF ORDNANCE RESEARCH DISTRIBUTION LIST

CONTRACT 447

Carnegie Institute of Technology

Commanding Officer
Office of Ordnance Research
Box CM, Duke Station
Durham, North Carolina 65

Chief of Ordnance
Department of the Army
Washington 25, D. C.
Attn: ORDTB-PS

Chief of Research & Development
Department of the Army
Washington 25, D. C.
Attn: Research Division

Commanding General
Aberdeen Proving Ground, Maryland
Attn: BRL, Tech Info Division

Commanding Officer
Detroit Arsenal
Center Line, Michigan

Commanding General
Frankford Arsenal
Bridge-Tacony Sts.
Philadelphia 37, Penna.
Attn: ORDBA-LC

Commanding Officer
Picatinny Arsenal
Dover, New Jersey
Attn: S. Feltman Ammo Lab

Commanding General
Redstone Arsenal
Huntsville, Alabama
Attn: ORDDW-MR

Commanding Officer
Rock Island Arsenal
Rock Island, Illinois
Attn: R & D. Division

Commanding Officer
Springfield Armory
Springfield 1, Mass.

Commanding Officer
Watervliet Arsenal
Watervliet, New York

Commanding Officer
Watertown Arsenal
Watertown 72, Mass.
Attn: OMRO

Commanding Officer
Watertown Arsenal
Watertown 72, Mass.
Attn: W. A. Laboratories

Commanding General
White Sands Proving Ground
Las Cruces, New Mexico
Attn: ORDBS-EML

Commanding Officer
Diamond Ordnance Fuze Laboratories
Washington 25, D. C.

Commanding General
Ordnance Weapons Command
Rock Island, Illinois
Attn: Research Branch

Contracting Officer
Pittsburgh Ordnance District
200 Fourth Ave.
Pittsburgh 22, Penna.

Chief of Ordnance
Department of the Army
Washington 25, D. C.
Attn: ORDGU-SE
For transmittal to:
Canadian Joint Staff
2001 Connecticut Ave., N. W.
Washington 25, D. C.

Office of Naval Research
Washington 25, D. C.
Attn: Code

The Director
Naval Research Laboratory
Washington 25, D. C.
Attn: Code 2021

Commanding Officer
Office of Naval Research
Branch Office
Navy, 100, FPO
New York, New York

Commander
U. S. Naval Proving Ground
Dahlgren, Virginia

U. S. Naval Ordnance Laboratory
White Oak, Silver Spring 19, Maryland
Attn: Library Division

Commanding Officer
U. S. Naval Powder Factory
Indian Head, Maryland

Chief, Bureau of Ordnance (AD3)
Department of the Navy
Washington 25, D. C.

Director
Air University Library
Maxwell Air Force Base
Alabama

Commander
Wright Air Development Center
Wright-Patterson Air Force Base
Ohio
Attn: WCRRO

Director
National Bureau of Standards
Washington 25, D. C.

NAC for Aeronautics
1512 H Street, NW
Washington 25, D. C.
Attn: Chief, Res. Coordination

Commanding General
Office of Scientific Research, ARDC
Bldg. T-1
Washington 25, D. C.
Attn: RDR
RDTOIL (Tech Library)

Technical Reports Library
SCEL, Evans Signal Corps Lab
Belmar, New Jersey

Commanding General
Signal Corps Engineering Lab
Fort Monmouth, New Jersey
Attn: Director of Research

Office of the Chief Signal Officer
Engineering & Technical Division
Engineering Control Branch
Room 2B273, Pentagon Building
Washington 25, D. C.
Attn: SIGGD

Commanding Officer
Engineering Res. & Dev. Laboratories
Fort Belvoir, Virginia

Commanding General
Army Ballistic Missile Agency
Redstone Arsenal
Huntsville, Alabama

Commanding General
Research & Engineering Command
Army Chemical Center, Maryland

Armed Services Tech Info Agency
Document Service Center
Knott Building, 4th & Main Sts.
Dayton 2, Ohio

U. S. Atomic Energy Commission
Document Library
19th & Constitution Ave., N.W.
Washington 25, D. C.

Director, Applied Physics Lab
Johns Hopkins University
8621 Georgia Ave.
Silver Spring 19, Maryland

I R I A
University of Michigan
Willow Run Research Center
Ypsilanti, Michigan
Attn: Shelia Coon Librarian

OFFICE OF ORDNANCE RESEARCH DISTRIBUTION LIST (continued)

Technical Information Service
P. O. Box 62
Oak Ridge, Tennessee
Attn: Reference Branch

Corona Laboratories
National Bureau of Standards
Corona, California

Officer in Charge
Jet Propulsion Laboratory
California Institute of Technology
Pasadena 3, California

US Army R & D Liaison Group
USAREUR, Northern Area Command
APO 757
New York, New York